

Phase diagram of the Wako-Saitô-Muñoz-Eaton β hairpin Model obtained from the partition function zeros

Julian Lee*

Department of Bioinformatics and Life Science, Soongsil University, Seoul 156-743

(Received 22 July 2014)

Abstract

I study the partition function zeros of the Wako-Saitô-Muñoz-Eaton (WSME) β hairpin model in the complex temperature plane. For various values of the entropy cost of disordering a bond, the zeros show clear locus corresponding to the folding transition. By extrapolating the locus to the real axis, transition temperature can be determined for various values of the entropy cost, leading to the phase diagram of the WSME β hairpin model.

PACS numbers: 64.60.De, 87.15.Cc, 64.60.an, 87.15.hp

Keywords: Partition function zeros, Protein folding, β -hairpin, Phase diagram

*Electronic address: jul@ssu.ac.kr; Fax: +82-2-824-4383

I. INTRODUCTION

Studying equilibrium and dynamic aspects of protein folding is of great interest in computational biophysics and statistical physics. Due to the computational costs of studying realistic all-atom model, simplified toy models such as lattice proteins have been a subject of extensive research[1–4].

A class of toy models called Go-like models bias the Hamiltonian toward its native structure[5–16]. The Go-like models are based on the assumption that the native structure of a protein contains much information on the folding process. In this work, I study the protein folding transition using Wako-Saitô-Muñoz-Eaton (WSME) model[5–8], a simple Go-like model that describes a protein as a one-dimensional Ising-like model with long-range interaction. In order to specify a model, a native structure of a protein of interest should be first given. I will restrict ourselves to a β -hairpin in this work.

I apply the partition function zeros method[17–38], a powerful method for studying a phase transition, to study the folding transition of a β -hairpin in the context of WSME model. The partition function zeros of WSME β -hairpin model in the complex temperature plane, called the Fisher zeros[18], were studied in my previous works[38]. The zeros for hairpin with even number of bonds could be obtained analytically in the limit when the entropic cost of ordering a bond is much larger than one[38]. The zeros for hairpin with even number of bonds, whose structure has more resemblance to the real β -hairpins, could be obtained only numerically, and the zeros for the β -hairpin with 15 bonds were described.

In this work extend the previous results, and obtain the Fisher zeros for a β -hairpins with 20 and 21 bonds, for various values of entropic costs of ordering a bond. Although the β -hairpin is a finite-size system, the folding transition temperature can be obtained by extrapolating the locus of zeros to the positive real axis. By plotting the transition temperature for various values of entropic cost, which plays the role of the magnetic field in a magnetic system, I can obtain the phase diagram of WSME β -hairpin model.

II. THE MODEL

WSME model describes a protein as a one-dimensional Ising model with long-range interactions[5–8]. A binary variable m_i is associated with each bond of a protein which

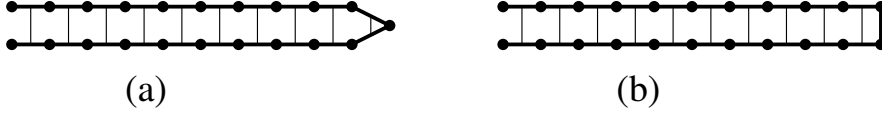


FIG. 1: The naive conformations of hairpin models with (a) $N = 20$ and (b) $N = 21$. The residues are denoted as filled circles, and the hydrogen bonds are shown as thin lines.

can take either one or zero depending on whether the bond is in the ordered or disordered state. Only the native contact contributes to the energy, and the contact can be formed if and only if all the intervening residues in the sequence are in the ordered states, so the Hamiltonian reads:

$$\mathcal{H}(\{m_k\}) = \sum_{i=1}^{N-1} \sum_{j=i+1}^N \epsilon_{ij} \Delta_{ij} \prod_{k=i}^j m_k \quad (1)$$

where N is the number of bonds, $\epsilon_{ij} < 0$ is the strength of contact between the i -th and the j -th bond, and $\Delta_{ij} = 1$ if the corresponding contact is formed in the native structure and zero otherwise. There is an entropy ΔS_i difference between an ordered and a disordered bond, and assuming that the ordered bond is described by a unique microstate, $\exp(\Delta S_i)$ the effective number of microstates of a disordered bond. The total number of microstates per bond is then $\Lambda_i \equiv \exp(\Delta S_i) + 1$.

The partition function can be written as:

$$Z = \sum_E \Omega(E; \{\Lambda_k\}) \exp(-\beta E) \quad (2)$$

where the density of states $\Omega(E; \{\Lambda_k\})$ is the number of microstates for given values of the energy E and the parameters $\{\Lambda_k\}$. Therefore, once the density of states is computed, the partition function is obtained from them in a straightforward manner.

The model is fully determined only after we specify the native structure of protein. In this work I will restrict myself to a simple β -hairpin with $N = 20$ and $N = 21$, where N is the number of peptide bonds, and the i -th bond makes contact with $N - i + 1$ -th bond. These hairpins are depicted in Figure 1. The conformation with odd value of $N = 2n + 1$ has the same number of native contacts n as that with $N = 2n$ except that it has an extra bond at the turn not participating in the native contact. The conformation with an extra bond at the turn is more realistic, but the model with an even value of N has an advantage that its zeros

can be obtained analytically in a certain limit. I will assume that the contacts are formed mainly due to the backbone hydrogen bond and that the sequence effects are negligible, so the interaction strengths are uniform: $\epsilon_{ij} = -\epsilon$. Then the energy of the conformation where j native contacts are broken is simply $E_j = j\epsilon$ where we shifted the energy by a constant amount so that the energy of the fully folded conformation is taken to be zero. I will also assume that the entropic cost of ordering each bond is the same throughout the chain: $\Lambda_i = \Lambda$. Then the density of states $\omega_j(\Lambda)$ for a given value j of broken native contacts then takes the form[38]:

$$\omega_j(\Lambda) = \begin{cases} 1 & (j = 0), \\ (\Lambda^2 - 1) \Lambda^{2j-2} & (1 \leq j \leq n \text{ for even } N) \text{ } (1 \leq j < n \text{ for odd } N) \\ (\Lambda^3 - 1) \Lambda^{2n-2} & (j = n \text{ for odd } N). \end{cases} \quad (3)$$

The partition function is then written as:

$$Z = \sum_j \omega_j(\Lambda) y^j \quad (4)$$

where $y \equiv \exp(-\beta\epsilon)$.

The zeros of the partition function Eq.(4) in the complex plane of y have been investigated for $N = 14$ and $N = 15$ [38]. In this work, I study the zeros of the partition function Z for $N = 20$ and $N = 21$, with the same value of $n = 10$, for various real values of Λ , to obtain the phase diagram of WSME β -hairpin model.

III. THE PARTITION FUNCTION ZEROS AND THE PHASE DIAGRAM

The partition function zeros for a given value of Λ are obtained as the ten roots to the polynomial equation:

$$Z(y) = \sum_{j=0}^{10} \omega_j(\Lambda) y^j = 0. \quad (5)$$

Since the coefficient are real, any complex roots form conjugate pairs. Note that the summation forms a geometric series for an even value of N if $\omega(0) = 1$ is replaced by $\omega(0) = 1 - 1/\Lambda^2$, which can be justified only for $\Lambda \gg 1$. In this case Eq.(5) can be solved analytically to yield zeros distributed on a circle [38]. Otherwise, the zeros have to be obtained numerically. It has been shown for $N = 14$ that the zeros lies close to the analytic locus even for Λ as small as 2.0 [38].

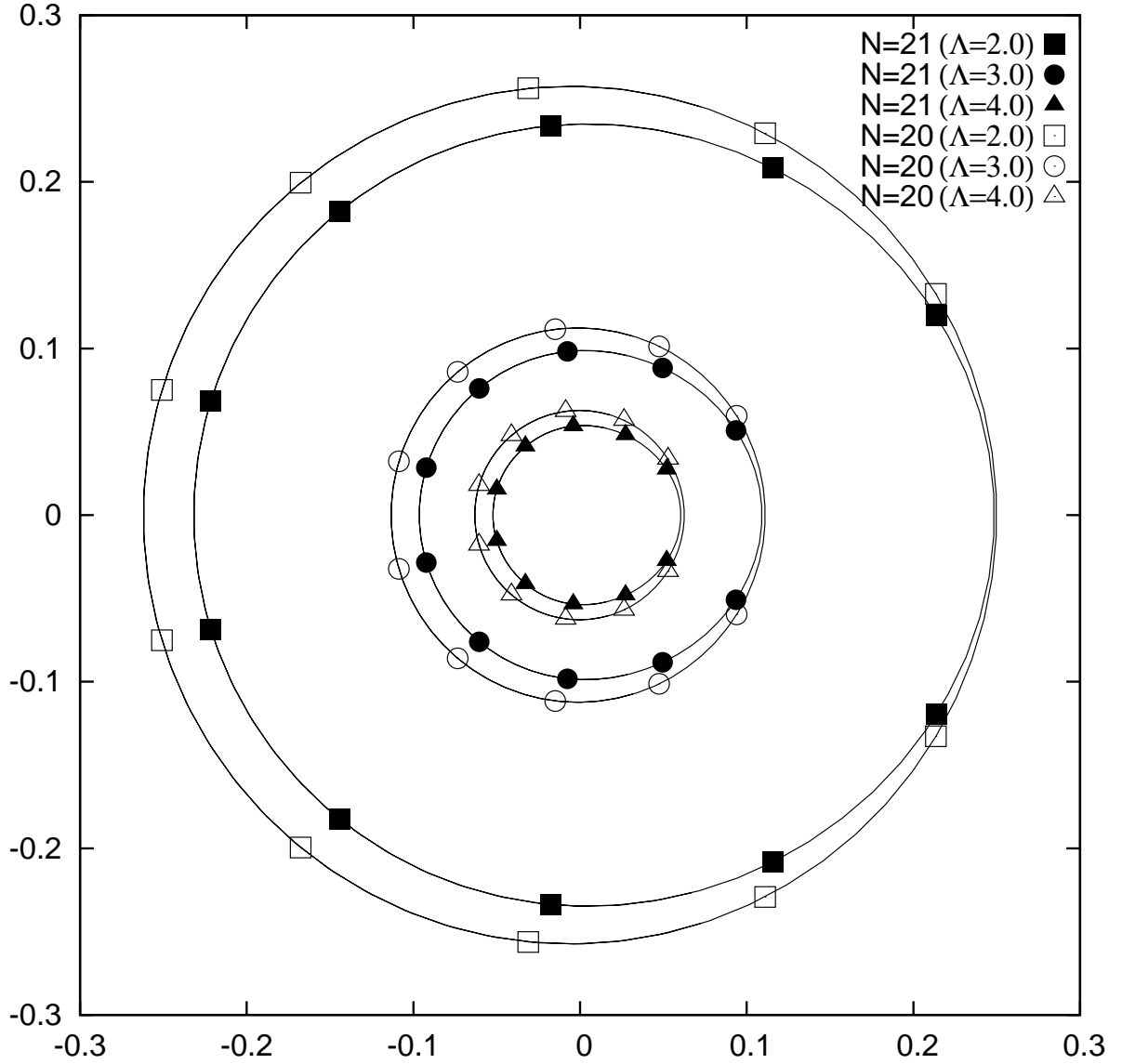


FIG. 2: The zeros for $N = 20$ (open symbols) and $N = 21$ (filled symbols). The curves obtained by extrapolating the zeros are also shown.

The zeros of the WSME β -hairpin for $N = 20$ and $N = 21$ are plotted in the Figure 2 for several representative values of Λ , which were computed using MATHEMATICA. The zeros form a clear visual locus that separate the complex plane into two regions. Although the β -hairpin is a system of finite size, a locus of PFZs can be used to *define* the phases of the system[20, 21]. Note that the physical range of $y \equiv \exp(-\beta\epsilon)$ is $0 \leq y \leq 1$, with $y = 0$ and $y = 1$ corresponding to $T = 0$ and $T = \infty$, respectively. Therefore, the region inside and outside the locus define the folded and unfolded phases respectively.

The loci of zeros closely resembles a circle for $N = 20$, but those for $N = 21$ have more skewed form, the distance of zero from the origin shorter near the negative real axis compared to those near positive real axis. As to be expected, the folded region shrinks in favor of unfolded region as Λ is increased. Since the zeros are exact, I extrapolated the locus to the positive real axis by obtaining a function expanded in terms of trigonometric function:

$$r(\theta) = \sum_{j=0}^4 A_j \cos(j\theta) \quad (6)$$

where r and θ are the polar coordinates in the complex plane of y :

$$y = r e^{i\theta} \quad (7)$$

A_j s are determined so that the zeros sit on the curve for $r(\theta)$. Since ten zeros form five conjugate pairs, there are only five independent zeros to be used, and only the even function of θ is allowed due to reflection symmetry with respect to the real axis. The method used here is similar to BST extrapolation[39–41], except that trigonometric function is used instead of rational function, since we are considering extrapolation with respect to an angle. The curve obtained by extrapolation is shown in Figure 2.

The transition temperatures obtained for various values of entropy per bond $S_b \equiv \ln(\Lambda)$ are plotted in Figure 3, leading to a phase diagram. The analytic formula of zeros for even values of N have obtained under the assumption of $\Lambda \gg 1$ leads to the folding temperature[38]

$$\frac{T_f}{\epsilon} = \frac{1}{2S_b} \quad (8)$$

which is also shown as a solid line in Figure 3. We see that the phase boundary defined from the exact zeros coincides quite well for $N = 20$, even for s as small as 0.31, corresponding to $\Lambda \simeq 1.37$, although the analytic formula was obtained under the assumption of $\Lambda \gg 1$. Even the phase boundary obtained for $N = 21$ coincides very well with the analytic curve, although we see a slight shift toward left for large values of S_b . This is to be expected, since the additional entropy cost for the extra bond at the turn will tend to favor the unfolded conformation. However, this effect is quite small, and as can be seen from the loci of zeros (Fig.2), the main effect of this extra turn is only to shift the loci near the negative real axis, transition temperature at the positive real axis being quite robust with respect to the existence of the extra bond at the turn.

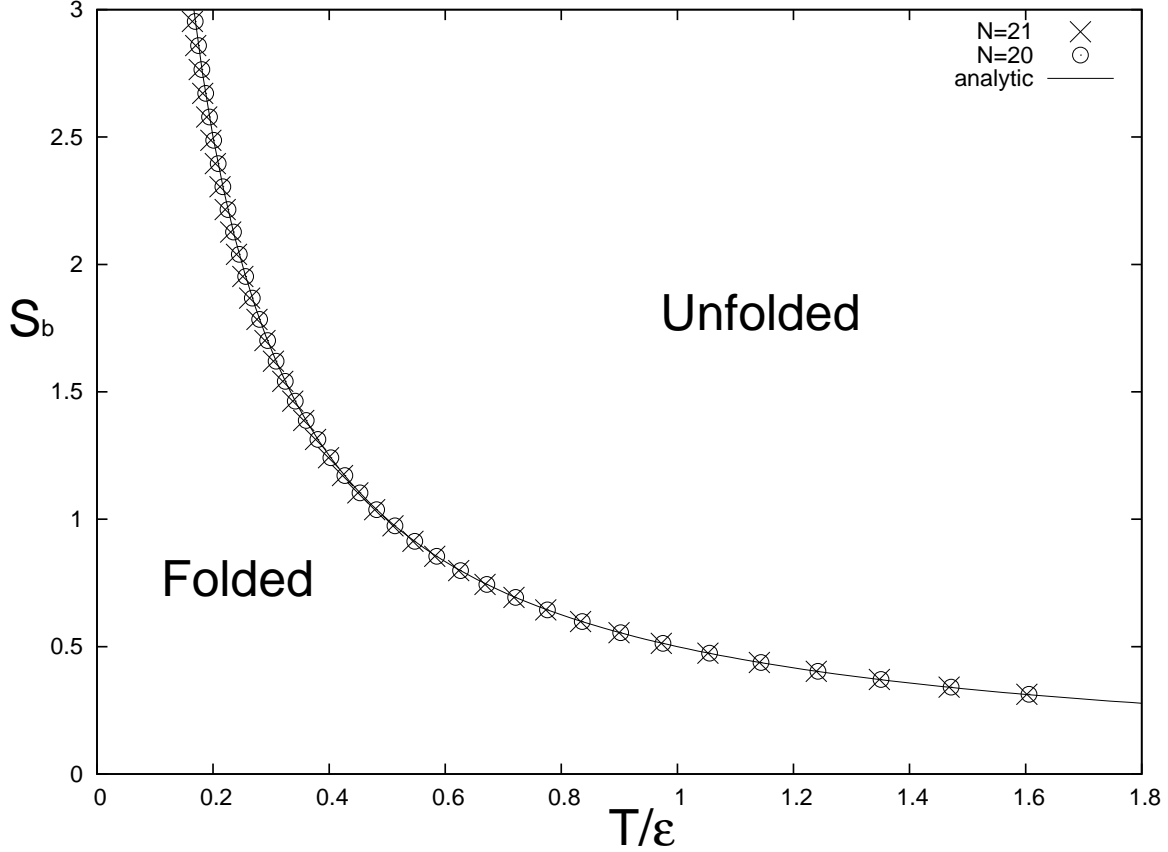


FIG. 3: The plots of the transition temperature for various values of $S_b = \ln(\Lambda)$, for $N = 20$ (circles) and $N = 21$ (crosses). The curve obtained by analytic formula are also shown.

To confirm that the low and high temperature regions correspond to folded and unfolded phases, I also drew the average fraction of unbroken native contacts $\langle Q \rangle$ as functions of temperature (Fig.4), for representative values of Λ . The transition temperature defined by the extrapolation of the zeros are shown as arrows, and we see indeed the average fraction of native contacts are about 0.5 at the transition point.

IV. CONCLUSIONS

In work, I studied the exact partition function of WSME β -hairpin model for $N = 20$ and $N = 21$. The model with $N = 21$ has the same number of native contacts as that with $N = 20$ except that it has an extra bond at the turn not participating in the native contact. I find that for a given value of the parameter Λ , the effective number of microstates per bond, the locus of the zeros for $N = 21$ is skewed near negative real axis compared to that

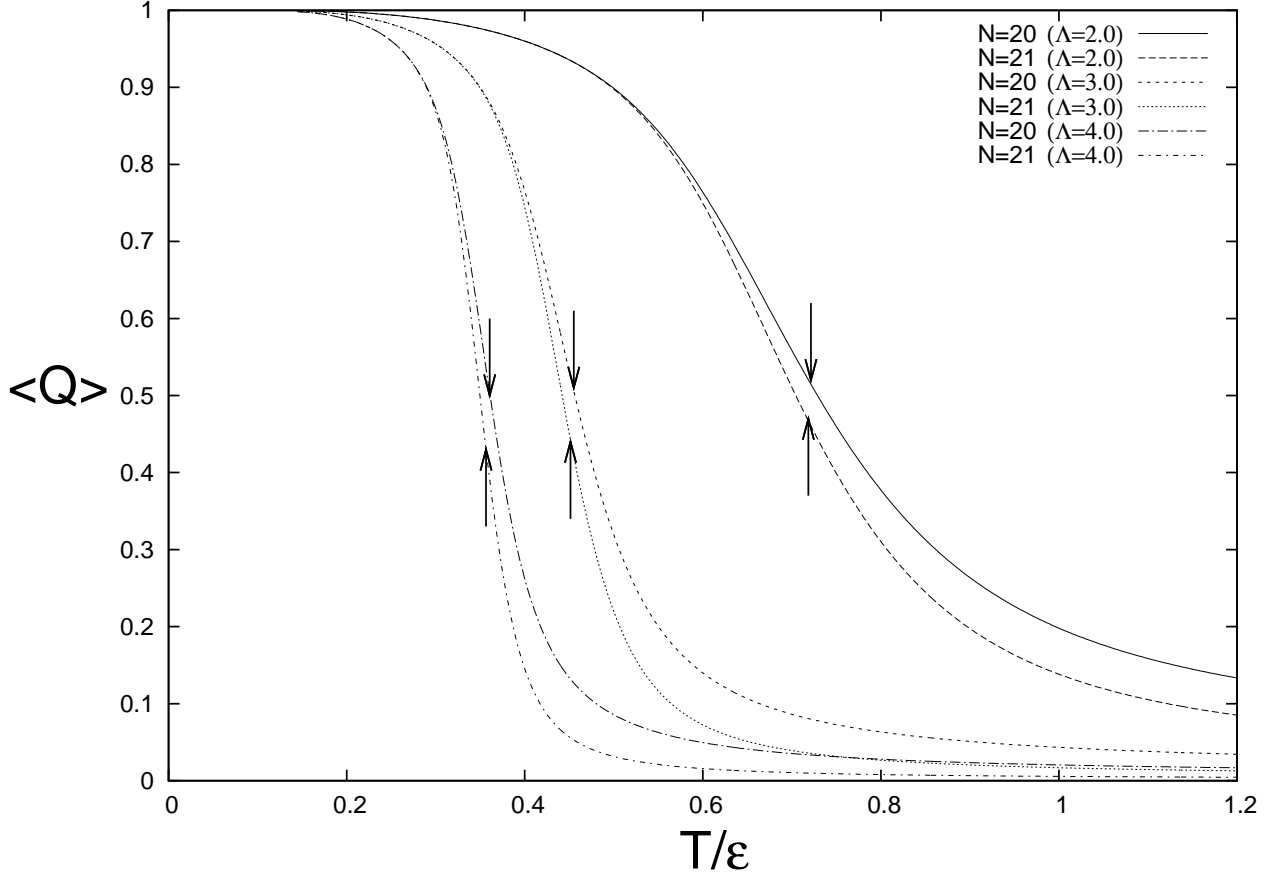


FIG. 4: The graph of the average fraction of unbroken native contacts $\langle Q \rangle$ as a function of temperature, for $\Lambda = 1, 2, 3$ and $N = 20, 21$. The transition temperatures defined by extrapolation of the zeros are indicated by the arrows.

of $N = 20$ which is closely approximated as a circle. However, zeros near the positive real axis are quite robust with respect to the existence of the extra bond at the turn. The phase diagram was obtained for various values of S_0 , the entropy cost for ordering a bond, and indeed the phase boundaries for $N = 21$ and $N = 20$ agree quite well with each other, as well as the analytic formula for even values of N obtained under the assumption of $\Lambda \gg 1$. I could observe a slight shift of the boundary toward the folded region for $N = 21$ and large values of S_0 , as to be expected. I also confirmed that roughly half of the native contacts are formed on average, at the transition temperature obtained by the extrapolation of the partition function zeros.

Acknowledgments

This study was supported by a Grant No. A120757 of the Korea Healthcare Technology R&D, Ministry of Health and Welfare, Republic of Korea.

- [1] K. F. Lau and K. A. Dill, *Macromolecules* **22**, 3986 (1989).
- [2] M. Bachmann and W. Janke, *Phys. Rev. Lett.* **91** 208105 (2003); *J. Chem. Phys.* **120**, 6779 (2004).
- [3] T. Wüst and D. P. Landau, *Phys. Rev. Lett.* **102**, 178101 (2009).
- [4] C.-N. Chen and C.-Y. Lin, *Physica A* **350**, 45 (2005).
- [5] H. Wako and N. Saitô, *J. Phys. Soc. Jpn.* **44**, 1931 (1978); **44**, 1939 (1978).
- [6] V. Muñoz, P. A. Thompson, J. Hofrichter, and W. A. Eaton, *Nature (London)* **390**, 196 (1997).
- [7] V. Muñoz et al., *Proc. Natl. Acad. Sci. U.S.A.* **95**, 5872 (1998).
- [8] V. Muñoz and W. A. Eaton, *Proc. Natl. Acad. Sci. U.S.A.* **96**, 11311 (1999).
- [9] E. Alm and D. Baker, *Proc. Natl. Acad. Sci. U.S.A.* **96** 11305 (1999).
- [10] O. V. Galzitskaya and A. V. Finkelstein, *Proc. Natl. Acad. Sci. U.S.A.* **96** 11299 (1999).
- [11] D. Baker, *Nature (London)* **405**, 29 (2000).
- [12] A. Flammini, J. R. Banavar, and A. Maritan, *Europhys. Lett.* **58**, 623 (2002).
- [13] P. Bruscolini and A. Pelizzola, *Phys. Rev. Lett.* **88**, 258101 (2002).
- [14] T. R. Weikl, M. Palassini, and K. Dill, *Protein Sci.* **13**, 822 (2004).
- [15] A. Imparato, A. Pelizzola, and M. Zamparo, *Phys. Rev. Lett.* **98** 148102 (2007).
- [16] K. Itoh and M. Sasai, *Proc. Natl. Acad. Sci. U.S.A.* **107**, 7775 (2010).
- [17] C. N. Yang and T. D. Lee, *Phys. Rev.* **87**, 404 (1952); **87**, 410 (1952).
- [18] M. E. Fisher, in *Lectures in Theoretical Physics*, Vol. 7c, edited by W. E. Brittin (University of Colorado Press, Boulder, 1965), p. 1.
- [19] C. Itzykson, R. B. Pearson, and J. B. Zuber, *Nucl. Phys. B* **220**, 415 (1983).
- [20] P. Borrmann, O. Mülken, and J. Harting, *Phys. Rev. Lett.* **84**, 3511 (2000); O. Mülken, P. Borrmann, J. Harting, and H. Stamerjohanns, *Phys. Rev. A* **64**, 013611 (2001); O. Mülken and P. Borrmann, *Phys. Rev. C* **63**, 024306 (2001).

- [21] W. Janke and R. Kenna, J. Stat. Phys. **102**, 1211 (2001); Comput. Phys. Commun. **147**, 443 (2002); Nucl. Phys. B **106**, 905 (2002).
- [22] N. A. Alves and U. H. E. Hansmann, Phys. Rev. Lett. **84**, 1836 (2000); Physica A **292**, 509 (2001); J. Chem. Phys. **117**, 2337 (2002); Y. Peng, U. H. E. Hansmann, and N. A. Alves, J. Chem. Phys. **118**, 2374 (2003).
- [23] W. Janke, D. A. Johnston, and R. Kenna, Nucl. Phys. B **682**, 618 (2004); R. Kenna, D. A. Johnston, and W. A. Janke, Phys. Rev. Lett. **96**, 115701 (2006); **97** 155702 (2006).
- [24] I. Bena, F. Coppex, M. Droz, and A. Lipowski, Phys. Rev. Lett. **91**, 160602 (2003).
- [25] J. Wang and W. Wang, J. Chem. Phys. **118**, 2952 (2003).
- [26] C.-N. Chen and C.-Y. Lin, Physica A **350**, 45 (2005).
- [27] I. Bena, M. Droz, and A. Lipowski, Int. J. Mod. Phys. B **19**, 4269 (2005) (and references therein).
- [28] Seung-Yeon Kim, Phys. Rev. Lett. **93**, 130604 (2004); Phys. Rev. E **70**, 016110 (2004); **71**, 017102 (2005); **74**, 011119 (2006); **81**, 031120 (2010); **82** 041107 (2010).
- [29] Jae Hwan Lee, Seung-Yeon Kim, and Julian Lee, J. Chem. Phys. **133**, 114106 (2010); **135**, 204102 (2011); Phys. Rev. E **86**, 011802 (2012).
- [30] P. F. Arndt, Phys. Rev. Lett. **84**, 814 (2000).
- [31] R. A. Blythe and M. R. Evans,
- [32] N. A. Alves, B. A. Berg, and S. Sanielevici, Phys. Rev. Lett. **64**, 3107 (1990).
- [33] Y. Iwasaki, K. Kanaya, and T. Yoshie, T. Hoshino and T. Shirakawa, Y. Oyanagi, S. Ichii, T. Kawai, Phys. Rev. D **46**, 4657 (1992)
- [34] A. Bazavov, B. A. Berg, D. Du, and Y. Meurice Phys. Rev. D **85**, 056010 (2012)
- [35] N. Borghini, R. S. Bhalerao, and J.-Y. Ollitrault, J. Phys. G: Nucl. Part. Phys. **30** (2004) S1213.
- [36] X. Zhu, M. Bleicher, and H. Stöcker, J. Phys. G: Nucl. Part. Phys. **32** (2006) 2181.
- [37] B. I. Abelev *et al.*, Phys. Rev. C, **77**, 054901 (2008).
- [38] Julian Lee, Phys. Rev. Lett. **110**, 248101 (2013); Phys. Rev. E **88**, 022710 (2013).
- [39] R. Bulirsch and J. Stoer, Numer. Math. **6**, 413 (1964).
- [40] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in C*, 2nd ed. Cambridge University Press, Cambridge, 1992, p.111.
- [41] J. L. Monroe, Phys. Rev. E **65**, 0661166 (2002).